Bayes の公式と Bayes 推定

樋口さぶろお

龍谷大学大学院理工学研究科数理情報学専攻

理論物理学特論 L11(2015-12-17 Thu)

最終更新: Time-stamp: "2015-12-19 Sat 13:40 JST hig"

今日の目標

1. Bayes の公式が説明できる
2. Bayes 推定できる

http://hig3.net

ここまで来たよ

1. Bayes の公式と Bayes 推定
    - 条件付き確率
    - ベイズの公式

2 変数の離散型確率変数の同時分布

6 枚のカードから無作為に 1 枚のカードを引く.

♡7　♡8　♡9 ◇8 ♠9 ♣9

同時分布

$X=$ 数, $Y=0$(赤札), 1(黒札) とすると (x,y) を得る確率

$P(X=x,Y=y)=f_{xy}^{\mathrm{XY}}$ は,

$$
f_{xy}^{\mathrm{XY}}=\begin{cases}\frac{1}{3} & ((x,y)=(8,0))\\ \frac{1}{6} & ((x,y)=(9,0))\\ \frac{1}{3} & ((x,y)=(9,1))\\ \frac{1}{6} & ((x,y)=(7,0))\\ 0 & (\text{他})\end{cases}
$$

2 変数以上のとき同時分布 結合分布 joint distribution という

表で書いた方がまし．ここでは，「他」は省略．

$y \backslash x$	7	8	9	計
0	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	
1	0	0	$\frac{1}{3}$	
計				

周辺分布

同時分布 f^{XY}_{xy} に対して，
X の周辺分布 $f^{\mathrm{X}}_{x} = \sum_y f^{\mathrm{XY}}_{xy}$.
Y の周辺分布 $f^{\mathrm{Y}}_{y} = \sum_x f^{\mathrm{XY}}_{XY}$.

要するに $\boxed{}$

連続型の周辺分布

$$f_{\mathrm{X}}(x) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}y, \quad f_{\mathrm{Y}}(y) = \int_{-\infty}^{+\infty} f(x,y)\,\mathrm{d}x$$

同時分布の母期待値

同時分布の母期待値

離散型　$$\mathrm{E}[\phi(X,Y)] = \sum_{x=-\infty}^{+\infty} \sum_{y=-\infty}^{+\infty} f(x,y) \cdot \phi(x,y)$$

連続型　$$\mathrm{E}[\phi(X,Y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) \cdot \phi(x,y) \mathrm{d}x\, \mathrm{d}y$$

同時確率と周辺確率

- 同時分布 $P(X=x,Y=y)$.
  - 意味 $X=x$ かつ $Y=y$
  - 性質
$$\sum_{x,y} P(X=x,Y=y)=1$$
.
- 周辺分布 $P(X=x), P(Y=y)$.
  - 定義
$$P(X=x)=\sum_{y} P(X=x,Y=y),$$
$$P(Y=y)=\sum_{x} P(X=x,Y=y)$$
.
  - 意味 Y は問わず $X=x$, X は問わず $Y=y$.
  - 性質
$$\sum_{x} P(X=x)=1, \qquad \sum_{y} P(Y=y)=1$$

条件付き確率 $P(X=x|Y=y)$, $P(Y=y|X=x)$

- 定義 (同時確率と周辺確率の比)

$$P(X=x|Y=y) = \frac{P(X=x,Y=y)}{P(Y=y)},$$
$$P(Y=y|X=x) = \frac{P(X=x,Y=y)}{P(X=x)}.$$

- 意味　条件 $Y=y$ のもとで $X=x$, 条件 $X=x$ のもとで $Y=y$.
- 性質 1 $\sum_x P(X=x|Y=y)=1, \sum_y P(Y=y|X=x)=1.$
- 性質 1' $\sum_y P(X=x|Y=y)\neq 1, \sum_x P(Y=y|X=x)\neq 1.$
- 性質 2 定義を通分して, 両辺に $\sum_y$ すると,

$$P(X=x|Y=y)P(Y=y) = P(X=x,Y=y)$$
$$P(X=x) = \sum_y P(X=x|Y=y)P(Y=y)$$

L11-Q1

Quiz(条件付き分布)

2 次元の離散型確率変数 (X, Y) を考える．同時分布 $P(X=x, Y=y) = f_{\mathrm{XY}}(x, y)$ は次の表で与えられる．

$y \backslash x$	2	3
3	2/12	1/12
7	5/12	4/12

1. 周辺分布 $P(X=x), P(Y=y)$ を求めよう．
2. 条件付き分布 $P(X=x|Y=3), P(Y=y|X=3)$ を求めよう．

L11-Q2

Quiz(ベイズの公式)

外見で区別できない, 品種 1(甘い) と品種 2(渋い) の柿がかごに入っている.
品種 1 は, 確率 0.95 で赤に, 確率 0.05 で黄色になる.
品種 2 は, 確率 0.125 で赤に, 確率 0.875 で黄色になる.
確率変数 X, Y を用いて, 品種 1(甘い) を $X = 1$, 品種 2(渋い) を $X = 2$, 赤いを $Y = 10$, 黄色いを $Y = 20$ と表現する.

1. 問題文から $P(Y = y|X = x)$ を読み取ろう.
2. かごの柿の $1/5$ が甘い柿であるとする. いま, 無作為に 1 個の柿を取りだしたところ, 赤い柿だった. ベイズの公式を使って, 取り出した赤い柿が甘い確率 $P(X = 1|Y = 10)$ を求めよう.
3. 仮にかごの柿の $1/5$ が渋い柿であるとする. いま, 無作為に 1 個の柿を取りだしたところ, 黄色い柿だった. ベイズの公式を使って, 取り出した黄色い柿が渋い確率を求めよう.

ここまで来たよ

1. Bayes の公式と Bayes 推定
   - 条件付き確率
   - ベイズの公式

ベイズの公式

$$P(X=x|Y=y) = \frac{P(Y=y|X=x)P(X=x)}{\sum_x P(Y=y|X=x)P(X=x)}.$$

$$P(Y=y|X=x) = \frac{P(X=x|Y=y)P(Y=y)}{\sum_y P(X=x|Y=y)P(Y=y)}.$$

$P(X=x|Y=y)$ を $P(Y=y|X=x)$ で書き表す式, およびその逆の式.

L11-Q3

Quiz(ベイズの公式)

確率変数 X は値 $x = 1, 2$, 確率変数 Y は値 $y = 10, 20$ をとり,

$$P(X = x) = \begin{cases} \frac{3}{4} & (x = 1) \\ \frac{1}{4} & (x = 2), \end{cases}$$

$$P(Y = y | X = 1) = \begin{cases} \frac{7}{10} & (y = 10) \\ \frac{3}{10} & (y = 20), \end{cases}$$

$$P(Y = y | X = 2) = \begin{cases} \frac{2}{5} & (y = 10) \\ \frac{3}{5} & (y = 20). \end{cases}$$

1. 同時確率を求めて表に書こう.
2. $P(X = x | Y = 10)$ を求めよう.

ベイズ的な考え方

$$\boxed{\text{事後確率 } P(X=x|Y=y)} \quad \longleftarrow \quad \boxed{\text{事前確率 } P(X=x)}$$

$$\uparrow$$

$$\boxed{\text{情報 } Y=y}$$

主観確率

ベイズの定理=ベイズの公式 (+ニュアンス?)

L11-Q4

Quiz(ベイズ推定)

抽選用の袋に何個かの色つきボールが入っている．ボールを割ると，中に当たり外れの記された紙が入っている．
当たりのボールのうち赤いボールが $\frac{1}{10}$，白いボールが $\frac{9}{10}$ である．
外れのボールのうち赤いボールが $\frac{7}{10}$，白いボールが $\frac{3}{10}$ である．
最初に，色は気にせず当たり外れだけ考えると，当たりの確率は $\frac{2}{10}$ くらいかなと思っていた (事前確率)．
無作為にボールを取り出したところ，赤いボールだった．このとき，外れである確率 (事後確率) はどれだけと思えるかを答えよう．
過程として同時確率の表を書くのを歓迎します．

連絡

- オフィスアワー月 4 木 6(1-502)